

R Notebook

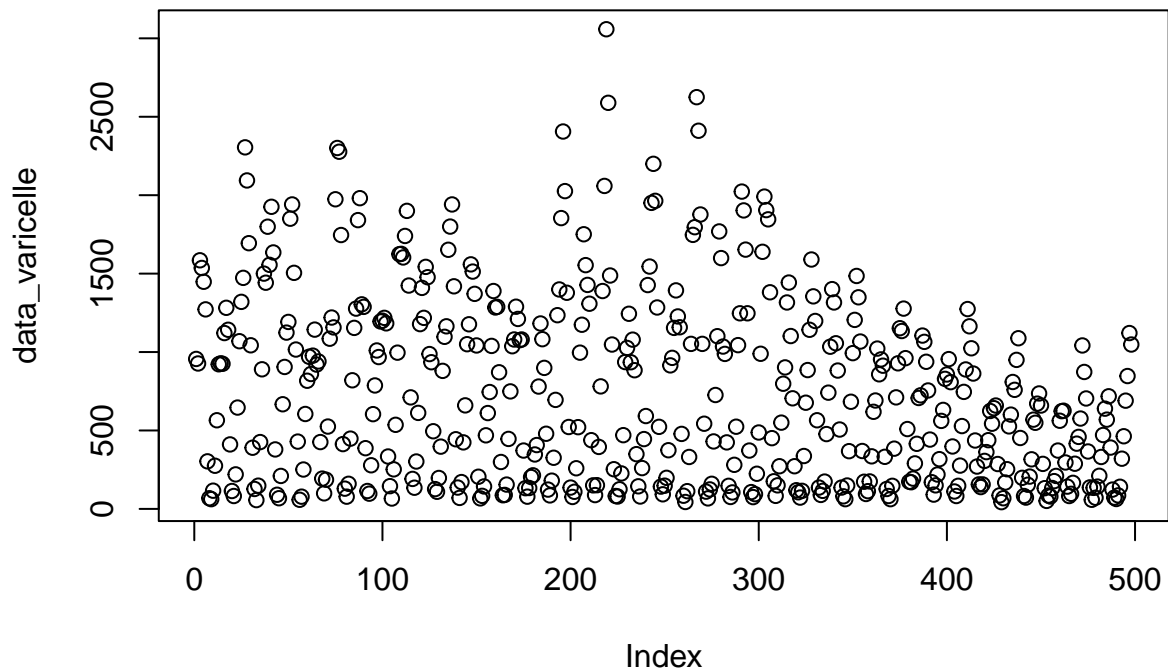
Mlamali SAID SALIMO

25 January, 2022, 11:37

```
data_varicelle = scan(file="data/varicelle.dat",skip=1)
head(data_varicelle)
```

```
## [1] 956 927 1585 1536 1448 1272
```

```
plot(data_varicelle)
```



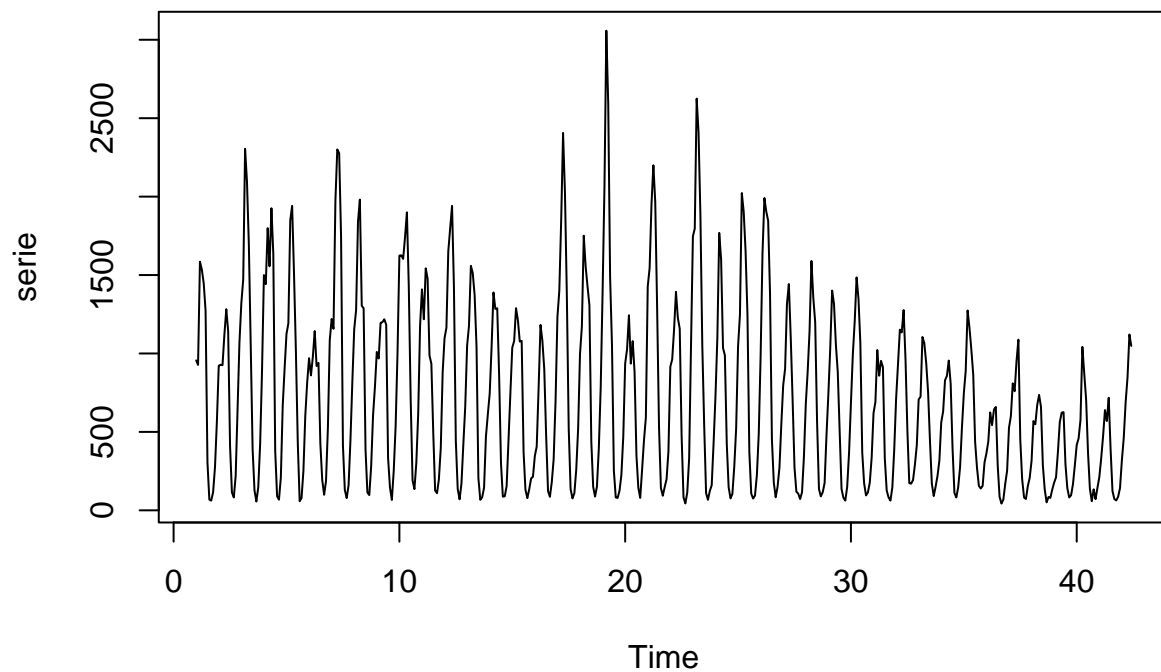
transforme les données en serie temporelles.

```
serie = ts(data_varicelle,frequency = 12) #frequence pour avoir les années en abscisse (les données me  
serie
```

##	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
## 1	956	927	1585	1536	1448	1272	303	68	62	116	275	565
## 2	922	928	925	1121	1282	1142	411	114	82	220	646	1069
## 3	1320	1473	2305	2094	1694	1043	390	127	56	148	426	890
## 4	1500	1442	1799	1556	1926	1635	379	90	68	210	667	905
## 5	1124	1192	1850	1941	1505	1016	429	58	78	251	605	817
## 6	970	860	977	1143	920	940	426	193	99	186	525	1085
## 7	1220	1157	1974	2301	2277	1746	413	129	78	160	448	820
## 8	1154	1277	1841	1981	1304	1288	387	114	97	278	604	787
## 9	1010	968	1195	1200	1218	1183	334	145	66	252	536	996
## 10	1624	1626	1603	1740	1900	1424	711	191	135	302	612	1178

```
## 11 1409 1218 1543 1477 987 935 495 126 109 197 397 880
## 12 1097 1164 1652 1800 1941 1419 444 136 70 171 424 660
## 13 1050 1177 1559 1513 1371 1042 205 67 83 143 469 611
## 14 745 1039 1389 1284 1288 871 299 87 89 155 446 749
## 15 1037 1080 1289 1211 1076 1080 372 132 78 133 203 214
## 16 347 407 780 1182 1082 899 479 123 86 180 326 695
## 17 1235 1399 1854 2406 2026 1378 522 136 76 109 259 521
## 18 996 1174 1751 1554 1428 1308 438 150 88 151 395 781
## 19 1389 2059 3058 2589 1488 1048 253 82 79 125 226 470
## 20 936 1026 1244 935 1079 884 349 144 79 260 445 592
## 21 1427 1545 1951 2200 1964 1284 523 142 93 148 198 374
## 22 915 963 1154 1393 1227 1158 478 84 44 113 331 1052
## 23 1747 1796 2625 2411 1877 1052 543 110 67 124 160 430
## 24 726 1101 1769 1599 1035 988 424 147 76 105 281 524
## 25 1044 1247 2023 1903 1653 1247 372 107 75 94 224 487
## 26 989 1639 1991 1905 1846 1381 451 176 83 150 272 550
## 27 798 902 1316 1443 1102 705 272 119 106 72 115 337
## 28 677 885 1142 1590 1355 1198 565 136 89 115 174 477
## 29 741 1034 1401 1316 1056 882 506 136 80 62 149 368
## 30 683 993 1205 1485 1349 1067 369 173 95 113 175 335
## 31 619 691 1022 858 953 913 332 127 82 62 147 384
## 32 711 928 1152 1134 1277 961 509 173 170 193 290 415
## 33 707 724 1105 1065 938 755 442 170 91 150 219 317
## 34 561 631 829 857 955 808 398 111 82 147 276 528
## 35 746 889 1274 1164 1024 863 436 270 156 139 156 306
## 36 362 438 624 543 642 659 286 86 43 68 168 253
## 37 526 601 809 759 950 1088 452 198 82 72 154 206
## 38 316 569 549 671 736 659 287 132 51 85 79 133
## 39 177 210 372 562 623 626 296 142 82 96 166 288
## 40 416 459 576 1042 873 704 366 137 58 134 71 142
## 41 211 331 471 639 569 718 391 123 72 63 86 141
## 42 320 463 690 847 1121 1048
```

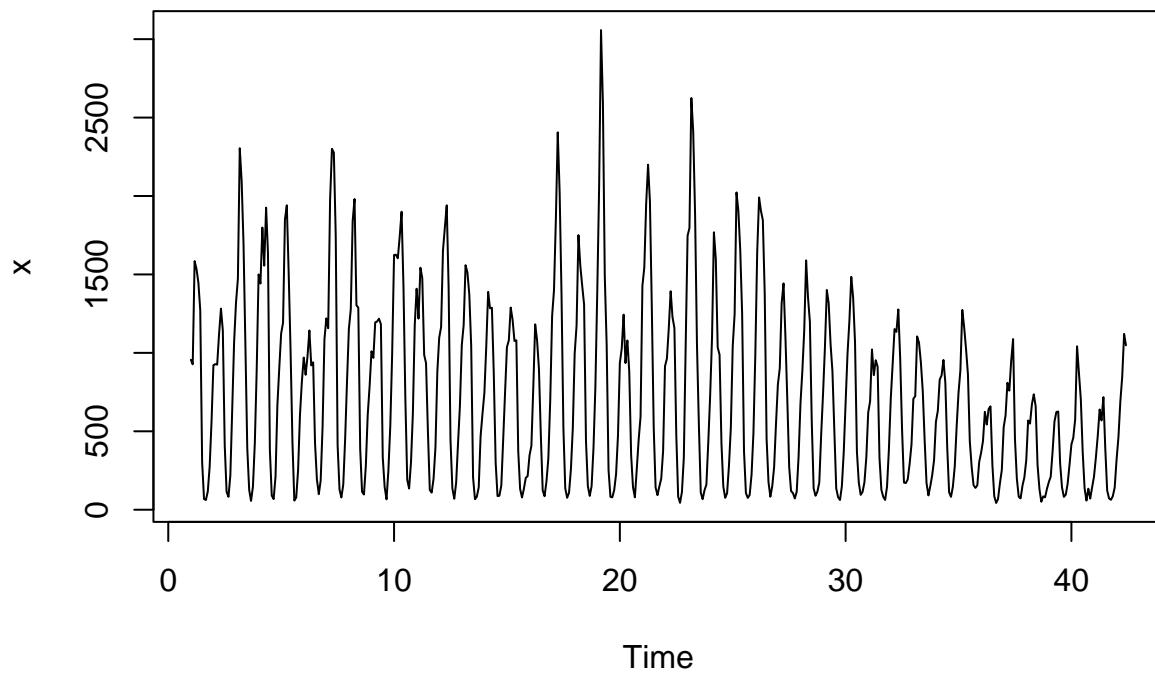
```
plot(serie) #dessiner le graphique
```



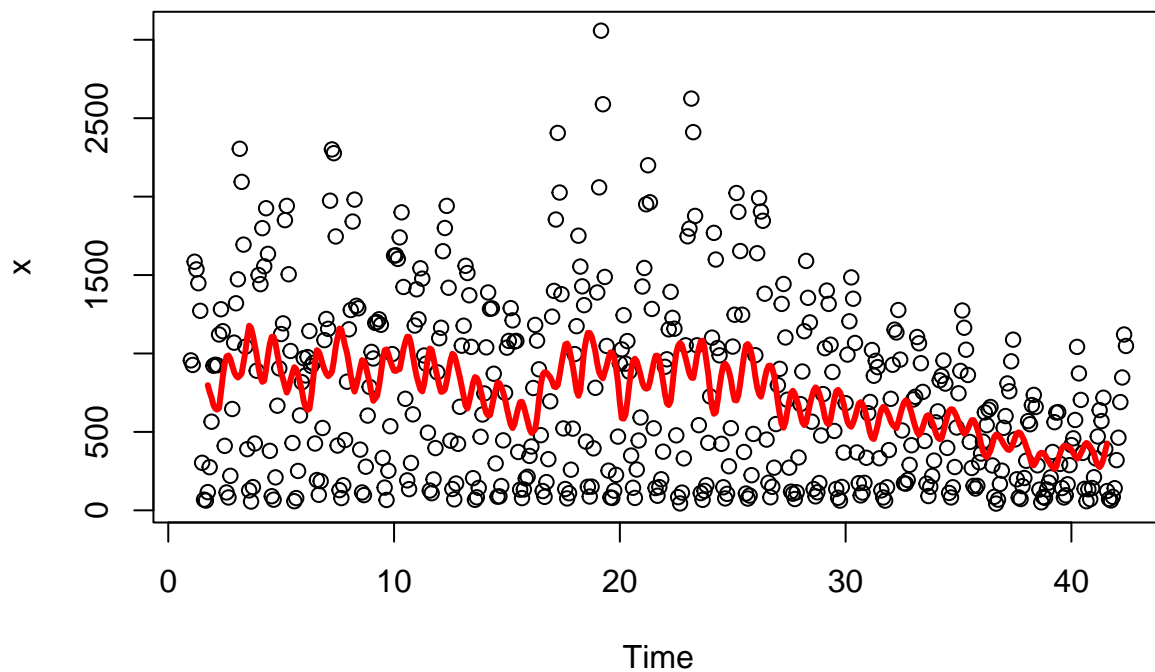
ne voit pas grand chose on peut lisser la courbe

si on

```
x = serie
plot(x)
```



```
plot(x,type='p')
k = 20
lines(filter(x,rep(1/k,k)), col='red', lwd=3)
```



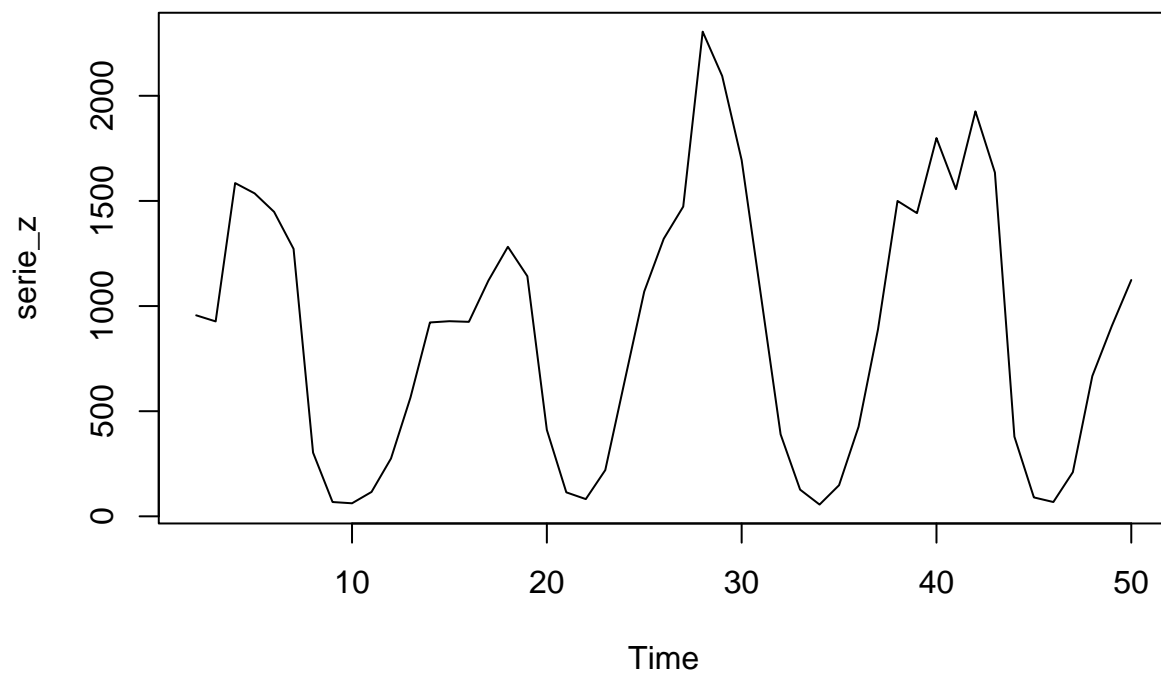
```
# prendre 20 valeur de x aléatoire
```

Question 2 - Analyser la série

On ne voit pas de tendance mais on voit un saisonnalité

Pour mieux voir la saisonnalité, on peut zoomer sur une durée plus petite. Les 50 premiers mois par exemple.

```
serie_z = ts(data_varicelle,2,50)
plot(serie_z)
```



Question 3 - calcul de la moyenne

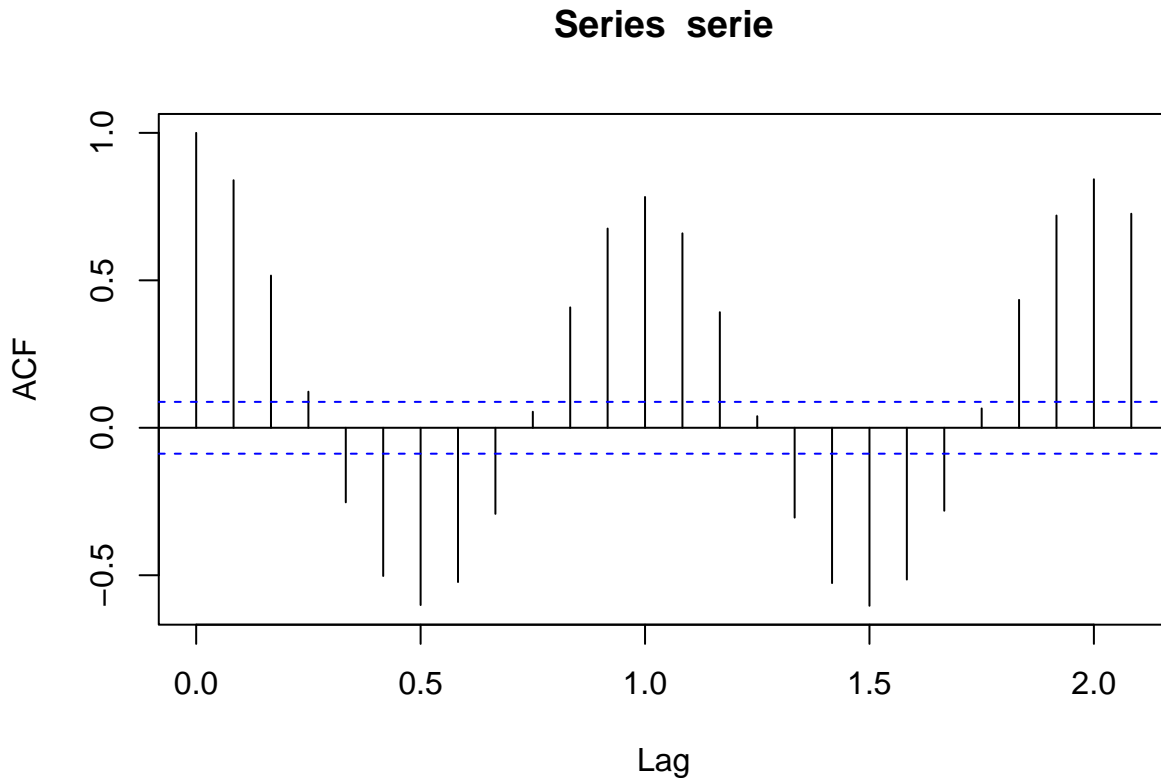
```
mean(serie)

## [1] 732.4076
```

Question 4 - ACF des 25 premières auto-corrélations

Def et objectif Fct de auto-corrélation :

```
#On rentre l'instruction
acf(serie,lag.max=25,type=c("correlation"))
```



on voit que ça converge, mais lentement. Tout ce qui est à l'intérieur de l'invalle on le considère statistiquement nulle (non significatif). Elle converge tjr vers 0 chakal la juste lentement

- on oberse une saisonnalité la saisonnalité oberseé graphiqmenet est confirmé par la fct de autocorréaltion.

On a fait avec 25 valeur et sa suffit largement

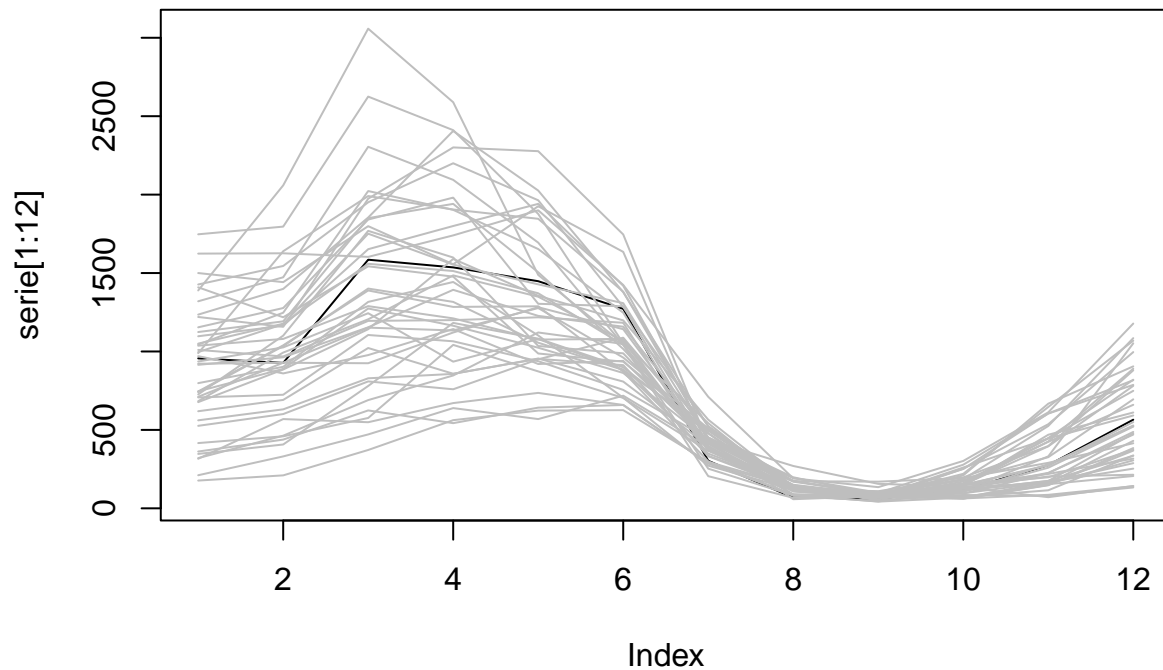
La fonction d'autocorrélation est périodique, ce qui indique une périodicité de la série temporelle.

La ligne pointillé bleu indique le niveau en dessous duquel la corrélation n'est plus statistiquement significative

Question 5 - Tracer sur un même graphique

Les évolution mesuelle de cas varicelles pour chaque année

```
plot(serie[1:12],type="l",ylim=c(min(serie),max(serie)))
for (i in 1 : 41) lines(serie[(1+12*i) : (12*(i+1))],col="grey")
```

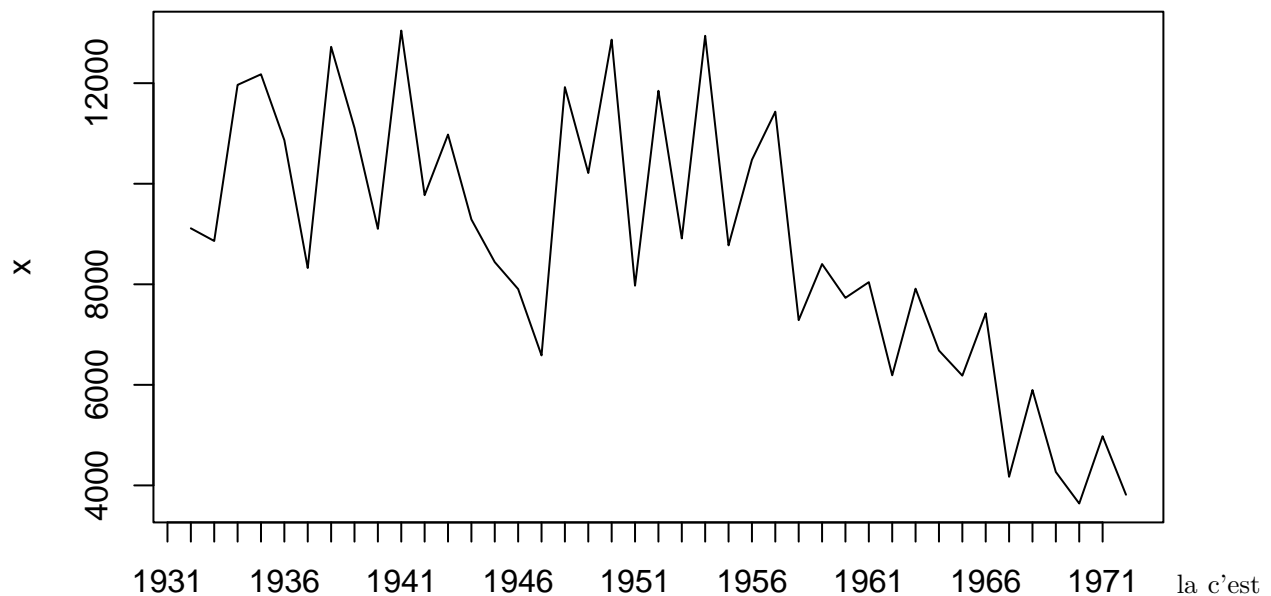


Question 6 : Tracer sur un graphique l..

```
x=rep(0,41)
for(i in 0 : 40) x[i+1] = sum(serie[(1+12*i) : (12*(i+1))])
x
```

```
## [1]  9113  8862 11966 12177 10866  8324 12723 11112  9103 13046  9773 10978
## [13]  9290  8441  7905  6586 11921 10214 12866  7973 11849  8912 12942  8775
## [25] 10476 11433  7287  8403  7731  8042  6190  7913  6683  6183  7423  4172
## [37]  5897  4267  3640  4978  3815
```

```
plot(x,type="l", xaxt='n', xlab='')
axis(1,at = 0 : 40,labels = 1931 : 1971)
```



évolution annuelles du nombre de cas de varicelle (tout à l'heure et l'évolution mensuelle pour chaque année)

Remarques : - a partir de 1955, la tendance de la diminution a été confirmée - ça c'est été maîtriser (la maladie) vers 1971

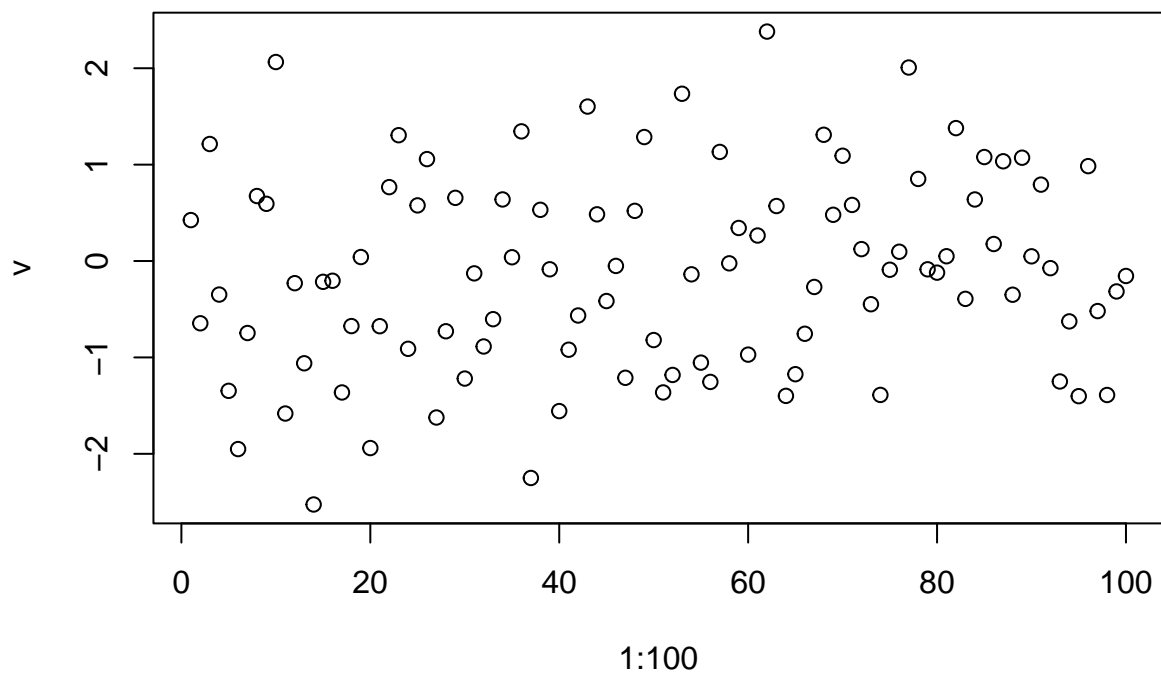
Exercice 2

1)

voir photo

2) Simuler un .

```
v = rnorm(100,0,1)
plot(1 : 100, v)
```

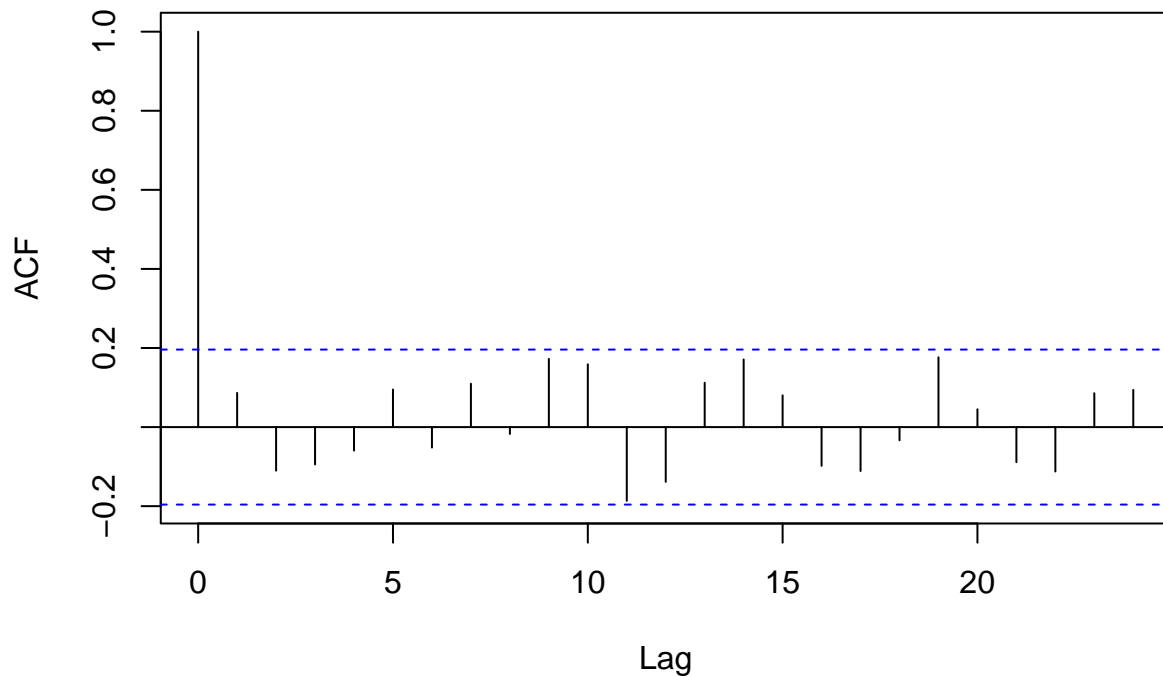


3)

traver la fct de auto

```
acf(v, lag.max = 24, type = c("correlation"))
```

Series v



- abs-
ciser : h, ordonnée $\rho(h)$ “normalement on devrai pas avoir plus de 5 valeur à l’exterieur de l’intervalle : psk nanani 5% param rnorm” “les termes sont corrélarder sauf ρ de 0, donc le graphique reflète bien la pertie théorique

y’a pas dévolution de la, sur le graphique tous les valeur de ρ de 1 sont tout statistiquement non signiicative
rappel : l’intervalle de confiance d’pende de n

4) Recommencer les deux question et oberser la variabilité

5) simuler la série temporelle

avant et un bruit blanc, mtn une serie avec tendance

```
v = rnorm(100,mean=0,sd=1)
t = 1:100

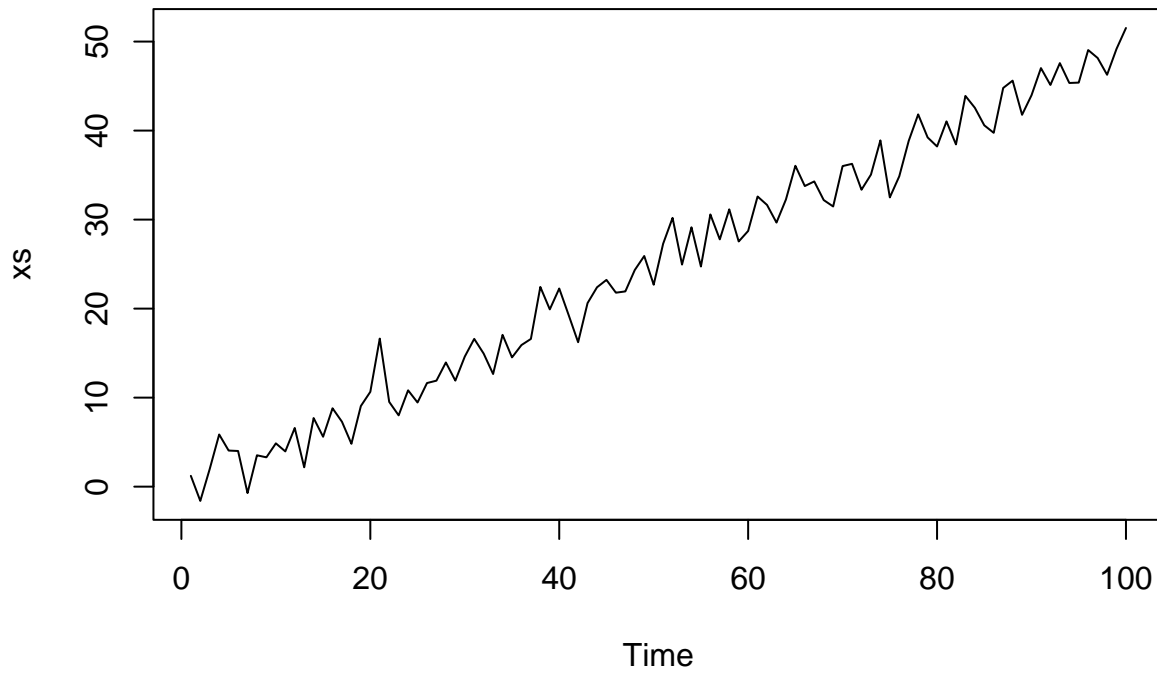
x = 0.5*t+2*v
xs = ts(x)
xs
```

```
## Time Series:
## Start = 1
## End = 100
## Frequency = 1
## [1]  1.2133059 -1.6022078  1.9899515  5.8582464  4.0539918  4.0056138
## [7] -0.7225218  3.5233432  3.2905608  4.8605782  3.9576691  6.5905056
## [13]  2.1750036  7.7060239  5.6049762  8.8052682  7.2881500  4.8124858
## [19]  9.0552000 10.6520499 16.6321219  9.5213722  8.0071141 10.8224489
## [25]  9.4523919 11.6385427 11.9046617 13.9519659 11.9108276 14.6106734
## [31] 16.6019102 14.9498879 12.6554001 17.0524327 14.5167420 15.9055101
## [37] 16.5939972 22.4348831 19.9132150 22.2594512 19.2685115 16.2237778
```

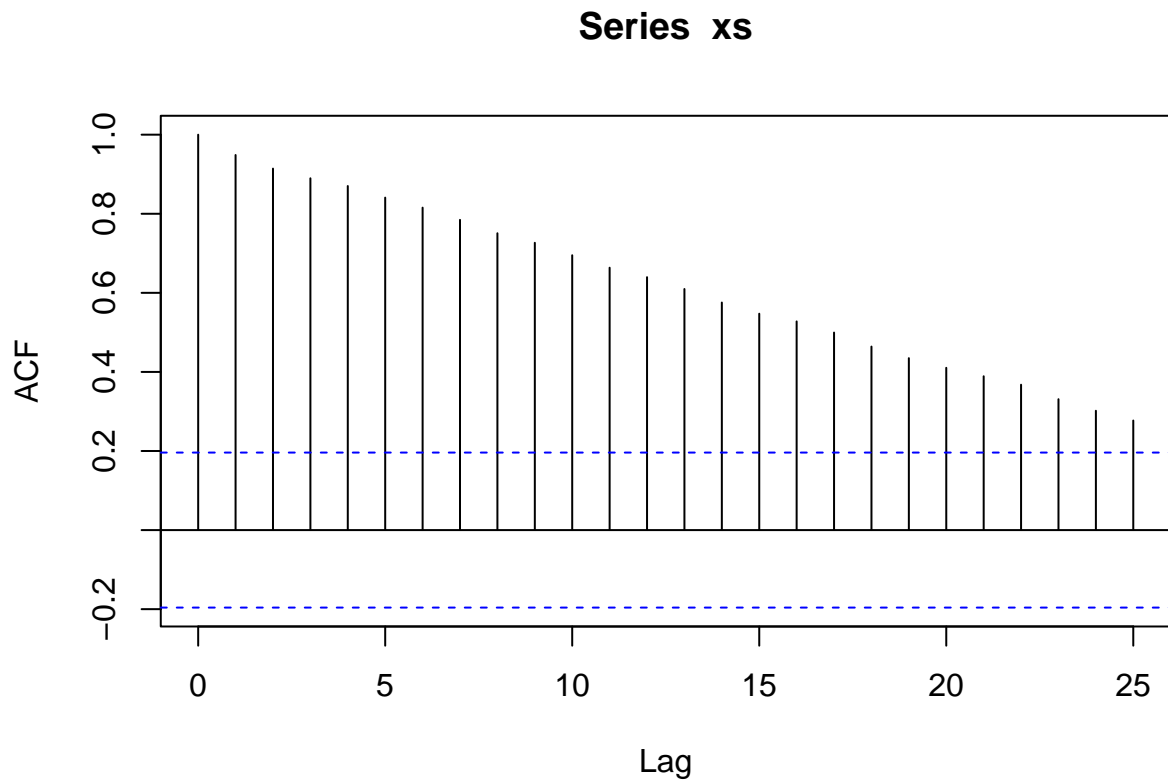


```
## [43] 20.6208052 22.3956707 23.2253240 21.7892489 21.9319808 24.3382097
## [49] 25.9118598 22.6831064 27.2842017 30.1945879 24.9484358 29.1351454
## [55] 24.7313932 30.5792973 27.7802642 31.1509081 27.5395239 28.7239078
## [61] 32.5927134 31.6463801 29.6635648 32.2594355 36.0466488 33.7696267
## [67] 34.2860146 32.1924020 31.4742603 36.0132046 36.2650717 33.3560218
## [73] 35.0530577 38.8971664 32.4939686 34.8642347 38.8493833 41.8196812
## [79] 39.2319024 38.2151357 41.0319144 38.4488064 43.8946673 42.5623936
## [85] 40.5901907 39.7480705 44.7878675 45.6119850 41.7727776 43.9650215
## [91] 47.0200082 45.1245964 47.5797067 45.3458868 45.3898506 49.0472032
## [97] 48.1538396 46.2713425 49.1787595 51.5237351
```

```
plot(xs)
```



```
acf(xs,lag.max = 25,type = c("correlation"))
```



Title :

Suite à tendance croissante en général

$$X_t = 5t + \text{Eps}_t$$

$$E(X_t) = 5t$$

la serie X_t n'est pas stationnaire, ya une tendance (la moyenne n'est pas constante). Cette tendance ça se voit graph de fct de autocorrélation : on observe une décroissance lente de quel converge vers 0 mais très lentement

Q7

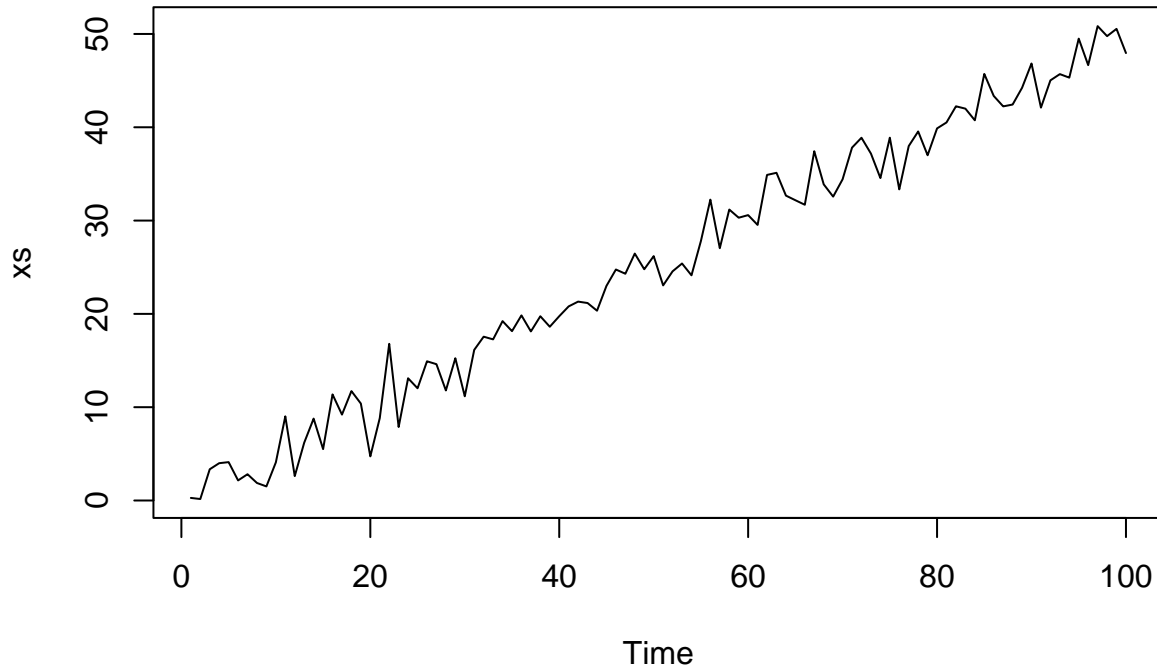
```
t = 1:100
v = rnorm(100,mean=0,sd=1)

x = 0.5*t+2*v
xs = ts(x)
xs
```

```
## Time Series:
## Start = 1
## End = 100
## Frequency = 1
## [1] 0.2779283 0.1585797 3.3468692 3.9972888 4.1126259 2.1510521
## [7] 2.8151296 1.8779228 1.5084234 4.0932542 9.0202973 2.6157133
## [13] 6.1917955 8.7758287 5.5085132 11.3690333 9.2047012 11.7258983
## [19] 10.3882163 4.7295343 8.8294486 16.7910053 7.8760238 13.1011067
## [25] 12.0261855 14.9195475 14.6089359 11.7939159 15.2442741 11.1682190
## [31] 16.1425848 17.5560176 17.2668388 19.2268042 18.1487310 19.8401146
## [37] 18.1172539 19.7468204 18.6233627 19.7497844 20.8033421 21.3147250
## [43] 21.1639415 20.3415914 22.9995594 24.7551988 24.3005124 26.4634426
```

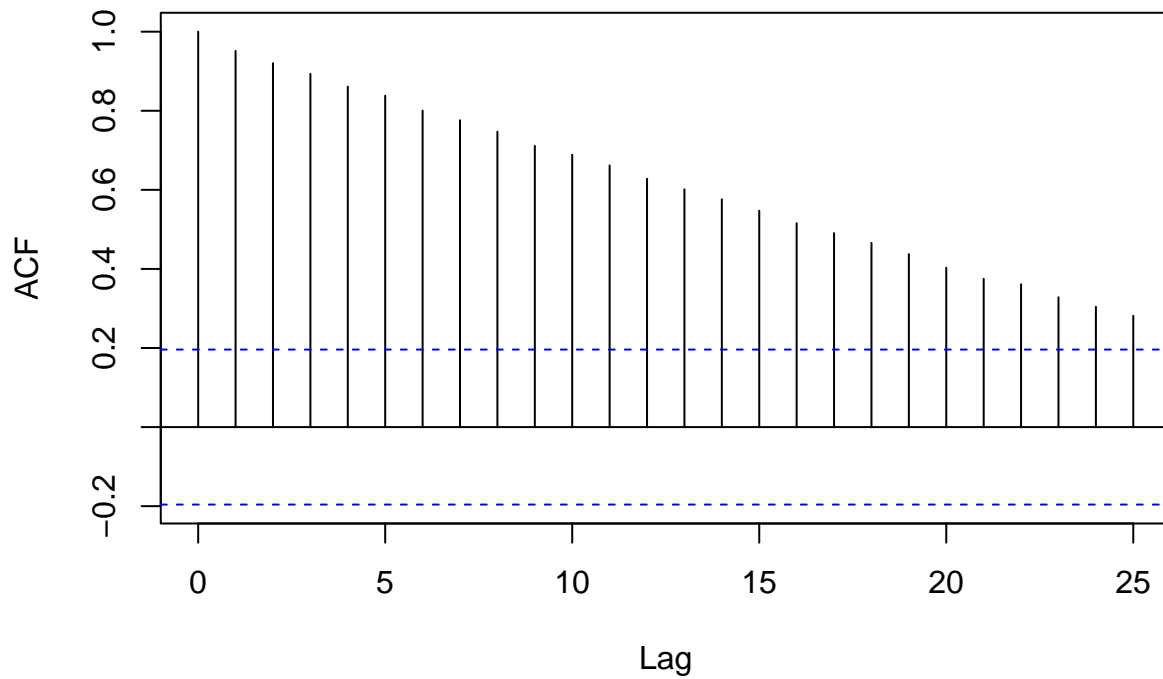
```
## [49] 24.7806769 26.1916966 23.0531353 24.5581839 25.3990460 24.1330851
## [55] 27.8131965 32.2421112 27.0379401 31.1840655 30.3058631 30.5867529
## [61] 29.5346753 34.8936325 35.1191381 32.6749056 32.1793284 31.6944673
## [67] 37.4310326 33.8801686 32.5652494 34.4228646 37.8194590 38.8779079
## [73] 37.1771620 34.5557417 38.8832832 33.3431360 37.9840076 39.5504097
## [79] 36.9985047 39.8779025 40.5144635 42.2457227 41.9946712 40.7443612
## [85] 45.7163935 43.3591455 42.2362540 42.4350015 44.2312569 46.8332344
## [91] 42.1047158 45.0262763 45.6970557 45.3197776 49.4976113 46.6600292
## [97] 50.8389340 49.7651036 50.5442654 47.9513634
```

```
plot(xs)
```



```
acf(xs,lag.max = 25,type = c("correlation"))
```

Series xs



Voire un petit souci. Normalement la saisonnalité doit se manifester là. pk elle se manifeste pas ? Parce quelle est absorbé par la tendance (il faut diminuer le coef sur le 0.5t)